Secure Heterogeneous Cloud Platform for Scientific Computing

To ensure the dynamism of modern business requires scalable and reconfigurable systems, however, the transformation of isolated and static corporate IT- resources is problematic.

Vladimir Zaborovsky, Alexey Lukashin
Department of Telematics
Saint-Petersburg State Polytechnical University
Saint-Petersburg, Russia
vlad@neva.ru lukash@neva.ru

***Abstract*— New technical systems and facilities are now using more accurate models that require high performance and large amounts of data to be processed. All these add new constraints on the effectiveness of configuration, scalability, and reliability of services. Data protection computation is used at various stages of the life cycle. In response to such demands, it is necessary to develop applications that can achieve high performance in heterogeneous computing infrastructure. Its components can function as in the modes of virtualization, and in the form of clusters, optimized for parallel calculations. Supercomputer Center «Polytechnic» is being created within the national research university, especially designed for high performance, scalability, heterogeneity and security resources for industrial applications and research. This paper proposes the way to use cloud services for hybrid high performance computing resources management and describes the implementation of hybrid cloud using heterogeneous computing resources, OpenStack platform, and stealth firewalls.**

Keywords-Cloud computing; security; heterogeneous platforms; firewalls; OpenStack

I. INTRODUCTION

Cloud providers, such as Amazon, Rackspace, Heroku, and Google may provide different services on the models of Infrastructure as a Service (IaaS), Platform as a Service (PaaS) or Software as a Service (SaaS), whose integration into a specific environment of industrial development is carried out by highly qualified engineers and IT- specialists. So far, actual challenge is to develop cloud services for scientific computing and computer aided engineering. These services have to provide human resources as well as computing environment. Existing engineering centers are being built today on a specially designed software and hardware platforms, which limits their performance and flexibility, or on the IaaS model that also does not allow you to efficiently solve a variety of engineering problems.

The center for Supercomputing Applications Platform "Polytechnic" was designed to solve a wide range of engineering tasks. It uses four types of systems combined to take into account the characteristics of different types of applications: system with globally addressable memory; hybrid cluster based on CPU and gpGPU; reconfigurable flow-computers; and cloud that span the computing systems in a single environment for shared storage of data and services to control access to resources. The use of such a heterogeneous computing environment has the following advantages:

- computing environment allows expanding the range of information services that allows to quickly and cost-effectively implement multidisciplinary projects;
- virtualization and heterogeneity provide scaling resources to ensure high performance of computation at all stages of the implementation of engineering projects;
- cloud architecture implements automatic configuration of hardware and software components, versioning of applications and monitoring the integrity of the computing environment;
- network services provide benefits of network centric approach in the implementation of complex engineering projects by geographically and logically distributed development teams and specialists;
- stealth security system implements a common policy in the field of information security.

There are different proposals and implementations for organizing scientific computing services using cloud services [1] [2]. In this paper we propose using IaaS OpenStack platform and security services for organizing heterogeneous engineering center.

Heterogeneous cloud platform is the basis of computing infrastructure for engineering centers; the principal difference from the classic data centers is to provide remote access not only for computing resources or applications, but also intelligent services that are implemented by teams of specialists in different areas working in outsourcing within the chosen corporate information security policy. Using the resources of modern cloud-based engineering centers it is possible to create equivalent social networks that bring together professionals and experts to perform multi-disciplinary engineering project including computations, verification of test results based on the use of different materials, virtual prototyping and data visualization of computation. The above-listed problems from the point of view at the computational algorithms can be combined into technological chains, which form a network of operations. Their implementation is provided within a heterogeneous

ISBN: 978-1-61208-330-8

cloud. The components of the platform (Figure 1), based on the OpenStack, include: IaaS cloud class segment, computing infrastructure within the cluster, the specialized high-performance hybrid system based on reconfigurable computing nodes.

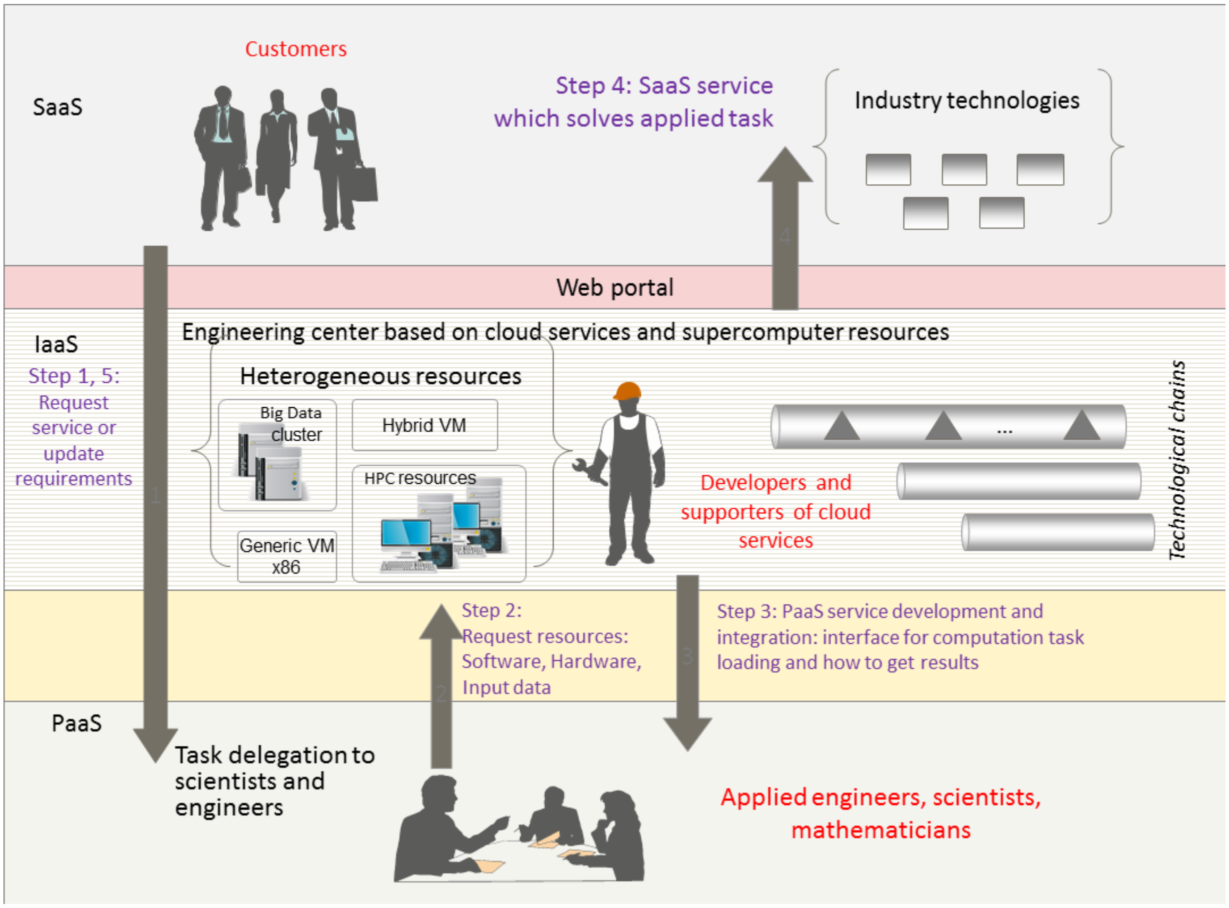

Figure 1. Functional scheme of a cloud platform with heterogeneous computing resources.

This paper is organized as follows. Section II covers security aspects of cloud platform. Section III covers methodology and technology overview of creating computation segments in cloud environment. The paper concludes with Section IV and presents future work on Section V.

II. Ensuring Security

Virtualization has changed the approach of deploying, managing, and using enterprise resources by providing new opportunities for consolidation and scalability of computational resources available to applications; however, this led to the emergence of new threats posed by the complexity and dynamic nature of the process of providing resources. These threats can lead to the formation of a cascade process of security violations, which are powerless to traditional data protection systems. The existing approaches like “Scan and Patch” do not work in a cloud environment — network scanners cannot track changes of resources configuration in real time. These approaches do not accurately identify the change in level of risk and take steps to block dynamically emerging threats.

To solve the problem of controlling access in the cloud, it is required to continuously monitor resources that cannot be provided without the automatically generating rules for filtering and firewall log files analysis. Information security management products in a dynamic cloud environment should include mechanisms that provide: total control over processes for deploying virtual machines; proactive scanning virtual machines for the presence of vulnerabilities and configuration errors; tracking the migration of virtual machines and system configuration to control access to resources. Therefore, within the center of the "Polytechnic" a series of measures are set out to improve information security resources, namely:

- Enhanced Control of virtual machines. Virtual machines as active components of the service are activated in the cloud application random moments, and Administrator cannot enter and exit virtual machine out of operation, until the security scanner checks the configuration and evaluate security risks.
- Automatic detection and scanning. Information security services are based on discovery of

vulnerabilities in the computing environment. This discovery is based on the current virtual machine configurations and on reports of potential threats that come from trusted sources, such as antivirus update servers.

- Migration of virtual machines. Proactive application migration is an effective method to control security.

users and services running at any given time. On the platforms of this type, there are rare situations when user needs a single virtual machine. Therefore, cloud services support the dynamic creation of secure networks with a set of preconfigured virtual machines. Secured networks are connected to the firewalls which are integrated with the distributed SDN switch Open vSwitch and OpenStack Neutron networking service.

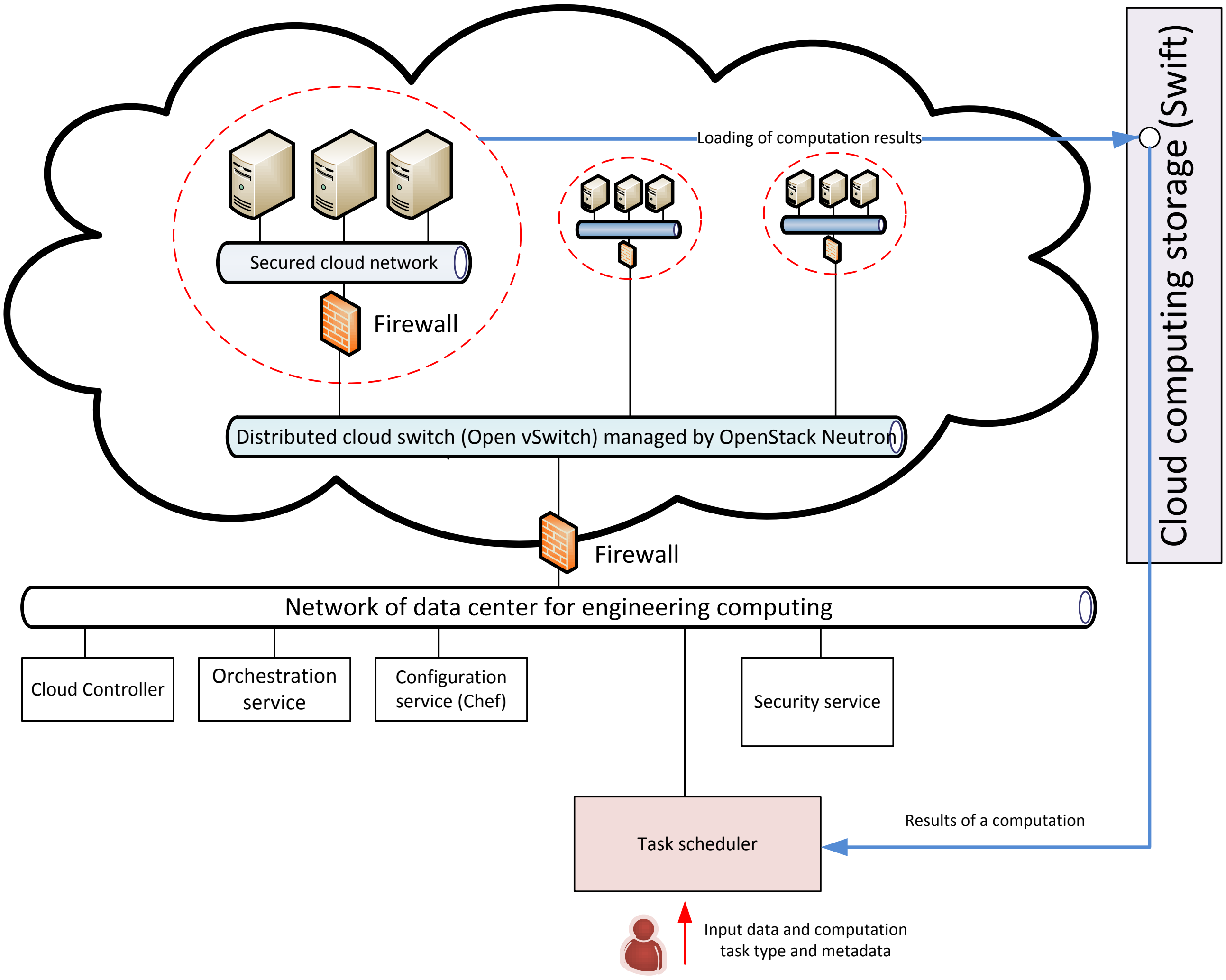
Loading of computation results
Secured cloud network
Firewall
Distributed cloud switch (Open vSwitch) managed by OpenStack Neutron
Firewall
Cloud computing storage (Swift)
Network of data center for engineering computing
Cloud Controller
Orchestration service
Configuration service (Chef)
Security service
Task scheduler
Results of a computation
Input data and computation task type and metadata

Figure 2. Components of cloud platform and general workflow.

In addition, computing center has the access control service for the cloud services protection. Its main features are support of dynamic infrastructure, scalability, and the ability to support security policies without reference to the composition of resources. The service is built on the technology of stealth traffic filtering and software defined networks (SDN) and provides the reconfiguration of access isolation system in accordance with the current state of the environment (in detail, the question of access control in cloud environments considered in [3] and [4]). Static platform segments applied the principle of "rent" under which the filter rules to access segments are formed only by

Firewalls which protect the dynamically generated cloud networks are created during computing segment initialization. An important feature of the firewall is its ability to function in the address less mode [5]. It allows implementing invisible protection of a cloud, and security system integration will not require the reconfiguration of a cloud network subsystem. The firewall acts as a virtual machine. The firewall of a network segment filters network traffic based on the rules that created the access policy service. Access policy in a cloud computing environment is based on the Role Based Access Control (RBAC) model of access control. This policy can be represented as a set of following attributes:

- user IDs, that are involved in the management of virtual machines and information services;
- privileges that are described in the form of permitted information services (privileges set rules for user access to services, it is possible to change the privileges for the user in the specified virtual machine filtering rules for your firewall, which allow access to a network service);
- set of roles that can be assigned to users;
- user sessions in a computing environments based on the network connections between subjects and objects.

Access policy is translated to firewall filtering rules according to computing environment state. This state can be represented by a set of IP addresses of computing resources, with assigned user labels. Label represents user which is responsible for computing resource. When a state of a computing environment changes, then it is necessary to generate a new set of filtering rules and reconfigure firewalls. For that, a method of the dynamic configuration rules is developed, which consist of substitution of the network address lookup in user-owners privileges for each virtual machine. This approach formed the rules of access to the services of the computational resource and of computing resource to services of other users.

It is required at least one virtual firewall in each virtualization server and one general bare-metal firewall for protecting cloud services from external threats. In a cloud-based system there is a dedicated management network separated from virtual machines, so this network is used for the information exchange between the components of the access control system and cloud services (Figure 2). OpenStack cloud platform is implemented by using service bus for communication between its components. Service bus is based on Advanced Message Queuing Protocol (AMQP) technology and RabbitMQ service [6]. Access control security service was integrated with OpenStack bus by subscribing its software components to events of OpenStack Compute service, which is managing the lifecycle of virtual machines, and OpenStack Neutron service, which is managing the lifecycle of cloud networks. When security service receives an event that a new virtual machine is starting, it generates and distributes filtering rules for the firewalls and generates rules for the virtual switch using OpenFlow technology which redirects traffic from virtual machine to the firewall. Firewall-based approach allows controlling traffic between instances which are connected to one virtual switch but belong to different users or security groups.

The proposed security service requires additional resources in the cloud. Traffic filtering costs make up about 10% of the virtualization server's resources [4].

III. PROTECTED SEGMENTS FOR ENGINEERING APPLICATIONS

For tasks which require heterogeneous computing resources, it is necessary to automate creation of the protected segments. We describe heterogeneous computing system as a set of logical computing resources. Such a segment must be applied to the specified security policy to permit the possibility of access to computing resources for the owner, but forbid access to these resources to other users. When the task is complete, the results must be loaded into the data warehouse, and the computing resources are freed. At the same time, it is essential to guarantee access to computing resources in simultaneous execution of multiple tasks.

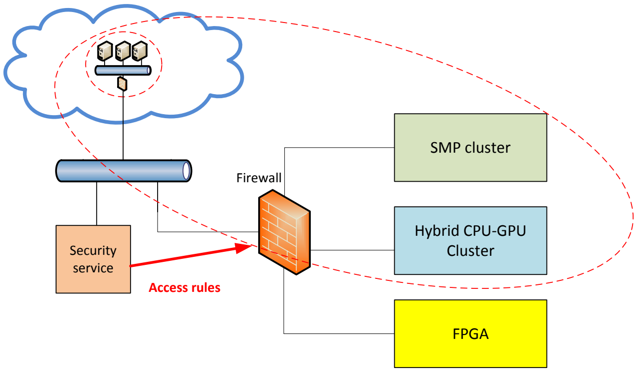

Figure 3. Reconfiguration of hybrid supercomputer center using firewalls.

We used OpenStack for creating groups of virtual machines in a cloud environment service. This service supports description of configurations in an Amazon Cloud Formation format that ensures compatibility with public services such as Amazon AWS. This service allows creating groups of virtual machines according to pattern, virtual networks, cloud-based routers and other components. The images of virtual machines contain a basic set of services. Any other application specific packages are installed using the automation services provided by Opscode Chef tool [7] that provides automated deployment of software configurations in virtual machines and bare-metal servers. When new computation segment is being created the security system spawns and configures virtual firewall which is filtering access to newly created network which serves computation. Dynamic network creation is supported by OpenStack Neutron services and distributed virtual switch Open vSwitch. After computing and receiving the results, the segment is removed, the cloud resources are released, and the results are uploaded to cloud storage and become available to the other consumers of the service. Every operation is automated: there are not any steps which need to involve human operations.

Reconfigurable segments of the cloud allow solving a wide range of scientific and technical tasks, among them: tasks that operate on large data sets based on the MapReduce technology; Bioinformatics tasks, including processing of genetic information in distributed systems; tasks of class CAD/CAE; calculation jobs not requiring high-speed networks. Tasks that cannot be solved in the cloud virtual machines (for example, requiring quick access to globally addressable memory and massively-parallel or streaming computations) are transferred to the dedicated hybrid clusters for high performance computing, computing infrastructure platform and equipped with an internal high speed communication bus, nodes-accelerators based on FPGA and GPU. Firewalls provide protection from unauthorized access to computing resources in a time of challenge and consolidation of heterogeneous segments (cloud and high-performance) computing resources into a single computation network, which components can communicate with each other, using the allowed protocols.

Built this way, infrastructure allows dynamically creating secure computing segments and thus provides an opportunity to organize a simultaneous solution of various tasks on a single set of hardware resources (Figure 3). The proposed solution implements reconfigurable federated cloud with one interface and multiple computation segments. A similar approach was used for organizing mobile cloud for intelligent transport systems and presented in [8].

IV. CONCLUSION

The proposed approach of organizing engineering center which is based on cloud services enables ability to reconfigure computing resources for different computation tasks. Integrated security services allow sharing computing resources between different users and clients. Reconfiguration of computing resources by using cloud firewalls is not a standard approach. It requires additional resources and makes platform more complex. From other side, it provides opportunity of reconfiguration of resources on network level. Stealth technology allows leaving applied software without modification. Dynamic computation segments creation service allows to effectively using IaaS resources on demand.

V. FUTURE WORK

The next step in our project is to integrate heterogeneous virtual machines in cloud platform. We are working on adding GPGPU devices like NVidia k20x and FPGA devices to virtual machines by using PCI pass-through capabilities in modern hypervisors like XEN and KVM. Recently released OpenStack Havana has support of PCI pass-through to virtual machines so we started evaluating how it works with heterogeneous compute devices.

REFERENCES

[1] D. Horat, E. Quevedo, A. Quesada-Arencibia, ”A hybrid cloud computing approach for intelligent processing and storage of scientific data”, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 8111 LNCS (PART 1), 2013, pp. 182 – 188

[2] J. Yimu, K. Zizhuo, P. Q. Yu, S. Yanpeng, K. Jiangbang, H. Wei, “A cloud computing service architecture of a parallel algorithm oriented to scientific computing with CUDA and monte carlo”, Cybernetics and Information Technologies 13 (SPECIALISSUE), 2013, pp. 153 – 166

[3] V. S. Zaborovsky, A. A. Lukashin, S. V. Kupreenko, and V. A. Mulukha, “Dynamic Access Control in Cloud Services. International Transactions on Systems Science and Applications”, ISSN 1751-1461, Vol. 7, No. 3/4, December 2011, pp. 264-277

[4] A. A. Lukashin and V. S. Zaborovsky. “Dynamic Access Control Using Virtual Multicore Firewalls”, The Fourth International Conference on Evolving Internet INTERNET 2012, ISBN: 978-1-61208-204-2, June 24-29 2012, Venice, Italy, pp. 37-43.

[5] A. Lukashin, V. Zaborovsky, S. Kupreenko, “Access isolation mechanism based on virtual connection management in cloud systems: How to secure cloud system using high perfomance virtual firewalls”, ICEIS 2011 - Proceedings of the 13th International Conference on Enterprise Information Systems 3 ISAS, 2011, pp. 371 – 375.

[6] S. Vinoski, “Advanced Message Queuing Protocol”, Internet Computing, IEEE, Volume: 10 , Issue: 6, 2006, pp. 87-89, doi 10.1109/MIC.2006.116.

[7] D. Spinellis, “Don't Install Software by Hand”, Software, IEEE, Volume: 29 , Issue: 4, 2012, pp. 86-87, doi 10.1109/MS.2012.85.

[8] V. .S. Zaborovskiy, A. A. Lukashin, S. G. Popov, A. V. Vostrov, “Adage mobile services for ITS infrastructure”, 2013 13th International Conference on ITS Telecommunications, ITST 2013, 2013, pp. 127 - 132